Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **460,** 4382–4396 (2016) doi:10.1093/mnras/stw1226
Advance Access publication 2016 May 26

Testing hydrodynamics schemes in galaxy disc simulations

C. G. Few,[1,2]★ C. Dobbs,[1] A. Pettitt[3] and L. Konstandin[1]

[1]*School of Physics, University of Exeter, Stocker Road, Exeter EX4 4QL, UK*
[2]*E.A. Milne Centre for Astrophysics, Department of Physics and Mathematics, University of Hull, Hull HU6 7RX, UK*
[3]*Department of Physics, Faculty of Science, Hokkaido University, Sapporo 060-0810, Japan*

Accepted 2016 May 19. Received 2016 May 12; in original form 2015 November 24

ABSTRACT

We examine how three fundamentally different numerical hydrodynamics codes follow the evolution of an isothermal galactic disc with an external spiral potential. We compare an adaptive mesh refinement code (RAMSES), a smoothed particle hydrodynamics code (SPHNG), and a volume-discretized mesh-less code (GIZMO). Using standard refinement criteria, we find that RAMSES produces a disc that is less vertically concentrated and does not reach such high densities as the SPHNG or GIZMO runs. The gas surface density in the spiral arms increases at a lower rate for the RAMSES simulations compared to the other codes. There is also a greater degree of substructure in the SPHNG and GIZMO runs and secondary spiral arms are more pronounced. By resolving the Jeans length with a greater number of grid cells, we achieve more similar results to the Lagrangian codes used in this study. Other alterations to the refinement scheme (adding extra levels of refinement and refining based on local density gradients) are less successful in reducing the disparity between RAMSES and SPHNG/GIZMO. Although more similar, SPHNG displays different density distributions and vertical mass profiles to all modes of GIZMO (including the smoothed particle hydrodynamics version). This suggests differences also arise which are not intrinsic to the particular method but rather due to its implementation. The discrepancies between codes (in particular, the densities reached in the spiral arms) could potentially result in differences in the locations and time-scales for gravitational collapse, and therefore impact star formation activity in more complex galaxy disc simulations.

Key words: hydrodynamics – methods: numerical – galaxies: evolution – galaxies: structure.

1 INTRODUCTION

It is well known that galactic dynamics play an important role in the formation of star-forming regions (e.g. Dobbs et al. 2014), with the gravitational potential of the spiral structure competing with hydrodynamical forces, large-scale differential rotation, and energetic feedback. The interplay of these processes is particularly difficult to study observationally and the hydrodynamical complexity means that a full understanding is analytically intractable. For these reasons, numerical simulations are a dominant tool for furthering our understanding of gas dynamics in a galactic context. However, the results can be quite different depending on the particular methodology for solving these equations. Hydrodynamics codes (which by now are extremely complex) may give conflicting results due to the respective strengths and weaknesses of the different implementations.

Code comparisons seek to quantify how different methodologies reproduce fluid flow and where weaknesses lie. These comparisons tend to concentrate on idealized test conditions (Agertz et al. 2007; Price 2008; Tasker et al. 2008; Hopkins 2015) which are easy to compare objectively, or on reproducing turbulent behaviour (Klessen, Heitsch & Mac Low 2000; Kitsionas et al. 2009; Price & Federrath 2010; Kritsuk et al. 2011), while others consider cosmological galaxy formation (Frenk et al. 1999; Pearce et al. 1999; O'Shea et al. 2005; Kereš et al. 2012; Scannapieco et al. 2012; Torrey et al. 2012). To date, almost no work (excepting Hopkins 2015) has compared the behaviour of isolated galaxy discs with different hydrodynamical codes

In this work, we compare the adaptive mesh refinement (AMR) code RAMSES (Teyssier 2002), the smoothed particle hydrodynamics (SPH) code SPHNG (Benz 1990), and the Lagrangian, mesh-less finite volume (MFM) code GIZMO (Hopkins 2015) in the context of a galactic disc with a rotating non-axisymmetric potential. Our purpose is to determine whether spiral galaxy and molecular cloud simulations with these codes are in concordance, and if not, what measures may be taken to achieve consistent results. This paper is organized as follows. In the remainder of Section 1, we briefly review existing code comparisons and relevant simulation methods. In Section 2, we describe the codes employed in this work, our initial conditions and the set of parameters we use. We analyse the growth of the spiral arms for the different codes and as a function of resolution in Section 3. We conclude with a discussion of our results in Section 4.

★ E-mail: c.gareth.few@googlemail.com.

1.1 Hydrodynamical methods

1.1.1 Smoothed particle hydrodynamics

SPH methods use the movement and concentration of particles to provide automatic refinement in dense areas. This means that low-density regions are more poorly resolved and the resolution of phenomena in areas where density gradients are steep are not guaranteed. Despite these issues, overdensities are almost always our regions of interest, and SPH methods are extremely useful.

We will not describe here the details of all SPH codes employed in this field but simply note works using them (e.g. Dobbs, Bonnell & Pringle 2006; Dobbs et al. 2008; Robertson & Kravtsov 2008; Saitoh et al. 2008; Grand, Kawata & Cropper 2012; Hopkins et al. 2012; Williamson & Thacker 2012; Dobbs & Pringle 2013; Mata-Chávez, Gómez & Puerari 2014; Williamson et al. 2014) and that typically resolution is given by particles numbers of $\sim 10^6$–10^7 and mass resolutions of the order of 10^3–10^4 $M_\odot$. Other than deliberate choices of cooling, star formation, and feedback etc., the key differences between these codes include how the smoothing length is determined, whether the SPH equations are used in pressure-energy or density-entropy form, and how artificial viscosity and conductivity are treated.

1.1.2 Grid-based hydrodynamics

The majority of other work in simulating isolated galactic discs employ some form of grid method. The simplest of these is a fixed Cartesian grid which may fit around a single spiral arm in the fashion of Kim & Ostriker (2002, 2006) or around the entire disc (e.g. Wada & Norman 2007; Wada 2008; Khoperskov et al. 2013) for which spatial resolution depends upon the size of disc that is simulated but can feasibly reach $\sim$7 pc. Of the works mentioned here only Khoperskov et al. (2013) simulates a disc with size comparable to the MW.

To reach greater resolution, grid simulations often use AMR (Berger & Oliger 1984; Berger & Colella 1989), whereby grids are subdivided to some appropriate resolution based on local hydrodynamical properties. The key strength of this approach is that it gives the user almost unlimited flexibility and control over which parts of the simulation volume are to be resolved and that large-scale effects can be incorporated into simulations where very fine spatial phenomena are to be studied. It should be noted though that this means that different works can employ different refinement schemes and parameters which influence the simulations.

One approach requires that the Jeans length be resolved by a minimum number of cells at each level. The minimum number is often set to the limit derived in Truelove et al. (1997), i.e. four grid cells per Jeans length. This refinement criterion is most important when considering self-gravitating gas and the Truelove et al. (1997) limit is intended to prevent numerical fragmentation, although resolving the Jeans length is also necessary to capture physical fragmentation of the gas. The Jeans length refinement criterion is employed in Tasker & Tan (2009), Renaud et al. (2013), Fujimoto et al. (2014), Petit et al. (2015) and Tasker, Wadsley & Pudritz (2015). Of particular note here is the work of Petit et al. (2015) in which the number of grid cells per Jeans length is increased from the typical 4 to a more rigorous value of 32.

Another approach, sometimes termed 'quasi-Lagrangian', resolves cells based on the local density such that the mass enclosed in a grid cell is roughly the same on each refinement level (e.g. Bournaud et al. 2010; Fujimoto et al. 2014; Agertz, Romeo & Grisdale 2015). One strength of this technique is that the mass per grid cell may also include stellar or dark matter mass which are usually gravitationally dominant. In this way, a stellar substructure can be well resolved for the gas phase even before the gas density increases. Note that even for isothermal runs in the absence of non-gaseous mass (such as presented in this work), a fixed mass threshold for refinement is not equivalent to a fixed Jeans length threshold.

The typical finest resolution for AMR runs ranges from 9 to 1.5 pc (considered sufficient to resolve the formation of the largest giant molecular clouds) with Renaud et al. (2013) achieving 0.05 pc resolution, albeit only for the last 50 Myr of the 780 Myr run. However, while oft quoted as 'the resolution', stating the minimum grid size is only slightly more informative than is the minimum smoothing length of an SPH simulation unless the phenomenon in which one is interested is entirely comprised of grid cells on that ultimate refinement level. Other refinement criteria are available, see Khokhlov (1998), but those described here cover the approaches currently employed in the simulation of galactic discs.

In addition to the criteria for refinement, and limits on the permitted levels, refinement usually also takes place in a cubic buffer around refined cells, this smoothes the transition between different levels and prevents a noisy mesh structure from forming: this buffer can vary in size and only ever increase resolution, refining cells that are not ordinarily qualified. Finally, some codes require derefinement criteria for the grid, but as RAMSES is not among them, we do not discuss this here.

1.1.3 Hybrid hydrodynamical methods

For a long time, the two dominant hydrodynamics schemes were SPH and grid (often AMR) methods, recently however new techniques have been developed. One new method, called Godunov-SPH or GSPH, involves a reformulation of SPH to introduce a Riemann solver that determines the force acting between each particle pair (Inutsuka 2002; Cha & Whitworth 2003). This does in principle grant the ability to resolve shocks in the absence of artificial viscosity and therefore avoid any side effects that occur as a consequence.

'Moving-mesh' techniques are a different approach to hydrodynamical codes which hybridize Lagrangian and Eulerian schemes. This technique is used in AREPO (Springel 2010). It evolves a finite-volume unstructured Voronoi mesh that moves with the fluid flow. This approach retains the ability of Eulerian codes to resolve shocks by employing a Riemann solver across each boundary between cells. Moving-mesh schemes are Galilean invariant and less noisy and less diffusive than SPH codes. AREPO is applied to the problem of an isolated galactic disc in Pakmor & Springel (2013) and Smith et al. (2014).

Another method that hybridizes SPH and grid codes is the so-called 'mesh-less' scheme that is detailed in Lanson & Vila (2008a,b), Gaburov & Nitadori (2011) and Hopkins (2015). In these codes, the particle ensemble is topologically similar to moving-mesh frameworks but for the lack of a sharply defined boundary between the resolution element domains. These codes are closer in their execution to SPH codes than moving-mesh techniques are but similarly employ a Riemann solver across the interfaces between resolution elements, enabling shock capturing. In this work, we consider one of these mesh-less codes, GIZMO (Hopkins 2015), in our comparisons.

1.2 Code comparisons

The differences between SPH and grid-based codes have been much discussed and we now briefly review the most relevant code comparison literature.

The first main type of comparison uses idealized tests that are simple setups for which analytical solutions exist such as the Sod shock tube or Sedov blast wave or instability tests (e.g. Kelvin-Helmholtz). Agertz et al. (2007) compare AMR and SPH using a similar number of resolution elements in the areas of interest. They found that all grid codes tested produced similar results for blob tests, Kelvin–Helmholtz and Rayleigh–Taylor instability tests and likewise, the SPH codes produced similar results as each other. Agertz et al. (2007) find that contact discontinuities and blob disintegration are less well reproduced by SPH codes, however Price (2008) counters that the inclusion of an artificial thermal conductivity term in SPH formulations reproduces these discontinuities. Likewise, the introduction of thermal diffusion improves instability resolution in SPH codes (Wadsley, Veeravalli & Couchman 2008). The inclusion of artificial conductivity/thermal diffusion is not relevant to this particular study because we are considering isothermal flows.

Idealized tests are presented in Tasker et al. (2008) with the conclusion that SPH codes excel at resolving the behaviour of dense objects while AMR codes are the preferred choice for voids (because they retain resolution in low-density regions) and shocks because of the ability to force resolution in areas of steep density contrast. This work also found that SPH codes can struggle to resolve fluid instabilities, while praising grid codes for being able to model multiphase fluids due to the sharp contrasts that can exist across grid boundaries. Despite these differences, the conclusion is that concordance between grid and particle codes is reached when there is one particle per grid cell in the region of interest.

The second main type of code comparison examines the turbulent properties of gas under the influence of some driving mechanism. Price & Federrath (2010) compare the ability of PHANTOM and FLASH to model supersonic driven turbulence and achieve similar results with comparable numbers of resolution elements (512^3). Despite similarities, PHANTOM is better at resolving dense structures, reaching densities at 128^3 resolution that FLASH only achieves at 512^3. The resolution of high-density regions is best achieved with SPH, but the grid resolves low-density structures better.

In Kitsionas et al. (2009), decaying turbulence is modelled with a variety of static grid and SPH codes. Grid codes are found to be less dissipative, but for the same number of resolution elements encouragingly similar results emerge. They conclude that resolution rather than the method primarily drives differences in turbulence simulations. Supersonic turbulence decay is also examined in Kritsuk et al. (2011) as a test of magnetodynamics codes. While the nature of the simulation set up is not completely relevant to this work we emphasize the result that even quite similar codes can generate different results due to minor code construction choices.

The third common framework for code comparison is galaxy formation in a cosmological context (Frenk et al. 1999; Pearce et al. 1999; Kereš et al. 2012; Scannapieco et al. 2012; Torrey et al. 2012). In works of this kind, insights into the accuracy of the hydrodynamical methods can be masked by dominant gravitational effects or by the method of analysis which usually focuses on derived physical properties not directly relevant to the scope of this study. Forthcoming comparisons in a cosmological context and with an isolated disc is expected under the AGORA project (Kim et al. 2014).

Finally, we mention Hopkins (2015) because in addition to a very detailed look at a number of idealized tests, the author also provides a comparative analysis of a cold Keplerian disc test and an isolated galactic disc, both of which are relevant to the work presented here. Hopkins (2015) shows the results of the viscous instability that affects SPH realizations due to shear viscosity in the cold Keplerian disc. Grid methods do not suffer from numerical viscosity in this context, but advection errors can cause rings to form in what should be a uniform disc. The work also presents two new methods in the form of a mesh-less lagrangrian code either with finite-mass (MFM) or MFV resolution elements, which perform well in this test and others. The analysis of isolated disc galaxy simulations is similar to what we present in this work. Those runs have lower particle mass resolution compared to ours but include star formation, stellar feedback, gas cooling and a live dark matter halo, central black hole, disc and bulge initialized in equilibrium. The finding here is that SPH runs are similar to one another and to the MFM realizations, however MFV transfers mass outwards due to an angular momentum advection error. It is particularly noteworthy that the critical flaw seen in SPH when applied to the Keplerian disc problem (viscous instability) does not manifest in this context because the pressure forces are much higher and the large stellar/gas mass ratio allows the stellar component to dominate and stabilize the gas disc.

The problem with determining the strengths and weaknesses of different methodologies using idealized tests is that the exact implementation used in those tests is often changed when the codes are applied to real problems. One example that stands out is that in idealized tests of AMR codes, different refinement criteria are applied to that which are used in production runs. In some cases, there is a justification for not applying the same refinement criteria, i.e. if the phenomenon being studied requires resolution based on the local density or if the ideal criteria would resolve an impractical amount of the volume which makes running the simulation intractable. This is the case in simulations of isolated galactic disc using AMR where resolution criteria are usually 'quasi-Lagrangian' or designed to resolve the Jeans length rather than resolving local gradients, as used for example when performing shock tube or blast wave tests.

2 METHOD

In this work, we compare three simulation codes: RAMSES, SPHNG, and GIZMO in a common framework. We simulate the evolution of isothermal gas from an initially uniform surface density disc under the influence of a disc galaxy gravitational potential with a fixed spiral perturbation. We now describe each of the codes employed in this work.

2.1 RAMSES

RAMSES (Teyssier 2002) is an AMR code in which gas dynamics are computed with a second-order unsplit Godunov scheme.[1] This scheme is inherently shock capturing with no need to invoke artificial viscosity. Time steps are advanced using a mid-point method with time centred fluxes at cell boundaries used to update the hydrodynamical variables. The time centred fluxes are determined using a second-order Godunov method (otherwise known as the Piecewise Linear Method). The duration of the time steps themselves is

[1] RAMSES is a publicly available code and can be found at http://www.ics.uzh.ch/teyssier/ramses/RAMSES.html

limited by a modified Courant–Friedrichs–Lewy (CFL) condition such that the time step is

$$\Delta t_{\rm CFL} = \frac{\Delta x_\ell}{\sum_{i=1}^{N_{\rm dim}}(|u_i| + c_{\rm s})} \frac{\sqrt{1 + 2C_{\rm CFL}\epsilon_{\rm GSR}} - 1}{\epsilon_{\rm GSR}}, \quad (1)$$

where Δx_ℓ is the linear extent of a grid cell at level ℓ with $N_{\rm dim}(=3)$ dimensional velocity $\boldsymbol{u}$ and sound speed $c_{\rm s}$. The right-hand part of equation (1) replaces the more traditional multiplication by the CFL factor, $C_{\rm CFL}$. This is changed here to vary with the gravitational strength ratio,

$$\epsilon_{\rm GSR} = \Delta x_\ell \frac{\sum_{i=1}^{N_{\rm dim}} |g_i|}{\left(\sum_{i=1}^{N_{\rm dim}}(|u_i| + c_{\rm s})\right)^2}, \quad (2)$$

where $\boldsymbol{g}$ is the gravitational acceleration experienced by each grid cell. The right-hand part of equation (1) is equal to $C_{\rm CFL}$ for $\epsilon_{\rm GSR} \rightarrow 0$ and is smaller (thus shortening the timestep) when the gravitational acceleration is large relative to the local gas velocity. The RAMSES runs presented here use $C_{\rm CFL}$=0.8. RAMSES supports different time steps for each level of the grid but we have enforced complete synchronicity so that the time step of coarse levels is identical to, and thus limited by, lower levels.

RAMSES is equipped with a number of solver options, the choices of which are rarely mentioned in literature. In this work, we follow Renaud et al. (2013) and use the acoustic Riemann solver with the MinMod slope limiter, however we also include a single run that uses the exact Riemann solver and the MonCen slope limiter. We set the mesh-smoothing parameter $n_{\rm expand}$=1.

In this work, we use different combinations of refinement criteria which are now described. The first criterion ensures that the Jeans length ($\lambda_{\rm J}$) is resolved by $N_{\rm J}$ cells. Each cell on level ℓ is marked for refinement if

$$\frac{\lambda_{\rm J}}{\Delta x_\ell} < N_{\rm J}. \quad (3)$$

An alternative criterion forces refinement where the local gradient in a variable q exceeds a fraction of the value of that variable. Refinement occurs if the following condition is satisfied,

$$C_{\rm q} < 2 \cdot {\rm MAX}\left[\left|\frac{q_{\rm i-1} - q_{\rm i}}{q_{\rm i-1} + q_{\rm i} + f_{\rm q}}\right|, \left|\frac{q_{\rm i} - q_{\rm i+1}}{q_{\rm i} + q_{\rm i+1} + f_{\rm q}}\right|\right], \quad (4)$$

where $i-1$ and $i+1$ are the cells that bound cell i in each dimension. The two user defined parameters here are the threshold $C_{\rm q}$ and a pseudo-floor value $f_{\rm q}$. In this work, we use only the gradient in density for grid refinement. In principle, the gradient in any hydrodynamical variable may be used, but as our simulations are isothermal, density and pressure criterion are degenerate and we have found that for our simulation set up, using velocity gradients as a criterion for refinement adds very little if density gradients are already being used.

One refinement scheme that we do not examine is the quasi-Lagrangian scheme whereby the mass enclosed in a given cell is compared to a threshold to determine if the cell is massive enough to warrant refinement. We do not examine this method because it is commonly used by combining the mass from the gas phase and stellar particles, the latter of which are not included in our simulations.

2.2 SPHNG

The SPH code used here (referred to as SPHNG) is a modified version of the code presented in Benz (1990). The density of a particle i is estimated through a weighted sum of the mass (m) of itself and its neighbours;

$$\rho_i = \sum_j m_j W(|\boldsymbol{x}_i - \boldsymbol{x}_j|, h_{ij}), \quad (5)$$

where the weighting function where W is the cubic spline kernel which is a function of the particle positions $\boldsymbol{x}$ and the mean smoothing length for each particle pair, h_{ij}. Particles are assigned variable smoothing lengths according to the local particle density (Price & Monaghan 2004). The smoothing length and density are solved iteratively via the Newton–Raphson method according to

$$h = \eta(m/\rho)^{1/N_{\rm dim}}, \quad (6)$$

where η=1.2 is equivalent to a typical number of ∼58 neighbours for each particle. Artificial viscosity is used to capture shocks following Monaghan & Lattanzio (1985) with parameters $\alpha_{\rm v}$=1 and $\beta_{\rm v}$=2 (after Monaghan 1992).

The code employs a second-order Runga–Kutta–Fehlberg integrator (Fehlberg 1969) to evolve the hydrodynamics equations. Particles are assigned individual timesteps, where the timestep of an individual particle is the minimum value from several limiters: a conventional CFL condition ($C_{\rm CFL}$=0.3)

$$\Delta t_{\rm CFL} = \frac{C_{\rm CFL} h}{c_{\rm s} + h|\nabla.\boldsymbol{v}| + 1.2(\alpha_{\rm v} c_{\rm s} + {\rm MIN}[0, \beta_{\rm v} h|\nabla.\boldsymbol{v}|)]}, \quad (7)$$

where $\boldsymbol{v}$ is the particle velocity, a force condition which limits the time-step depending on the net acceleration $\boldsymbol{a}$ on a particle,

$$\Delta t_{\rm a} = C_{\rm CFL}\sqrt{\frac{h}{|\boldsymbol{a}|}} \quad (8)$$

and a third requirement that changes in a particle's velocity, acceleration and smoothing length do not exceed a given tolerance, full details are given in Bate, Bonnell & Price (1995).

2.3 GIZMO

GIZMO (Hopkins 2015) is a hydrodynamics code, based on the SPH code GADGET-3, designed to accommodate the benefits of both grid and particle based hydrodynamics schemes.[2] The method is based on the works of Lanson & Vila (2008a,b) and Gaburov & Nitadori (2011), and has some common features with moving-mesh codes such as AREPO (Springel 2010). GIZMO uses a Lagrangian-like formulation, where the volume is discretized using a weighting function. The weighting function is similar to the kernel in SPH though in contrast, the kernel gradients play no role in the equations of motion as they do in SPH. The discretization of the fluid is defined by an ensemble of particles that trace the motion of the cells. Shocks are captured with a Riemann solver, eliminating the need for artificial viscosity (similar to the approach of Godunov SPH codes). There are two new methods available in GIZMO (as well as two different versions of SPH with a number of viscosity switches); MFV and MFM methods. The methods differ in whether the particles/cells are allowed to experience a mass flux between their neighbours (MFV) or whether their masses are fixed (MFM). The methods appear to differ only slightly in the test problems shown in Hopkins (2015), especially relative to the differences seen when compared to pure grid or SPH methods.

[2] GIZMO is a publicly available code and can be found at http://www.tapir.caltech.edu/phopkins/Site/GIZMO.html

The time integration scheme used in GIZMO is a second-order leapfrog integrator very similar to AREPO and GADGET2 (Springel 2010) and is described in detail in the appendix G of Hopkins (2015). Local time-steps are used in all modes so that for each particle the timestep is

$$\Delta t_{\rm CFL} = 2C_{\rm CFL}\frac{h}{|v_{\rm sig}|}, \quad (9)$$

where h is the kernel smoothing length and $|v_{\rm sig}|$ is the signal velocity (Whitehurst 1995; Monaghan 1997; Hopkins 2015). For the GIZMO runs, we use a CFL factor of $C_{\rm CFL}$=0.1, however note that the CFL factor is used differently by each of the codes and in particular the way that GIZMO and SPHNG employ this value is not directly comparable with the CFL factor used in RAMSES due to the different way in which resolution is defined for each framework. Time-steps in GIZMO are also limited to prevent spurious events of particle interpenetration when neighbouring particles have very different in time-steps (Saitoh & Makino 2009; Durier & Dalla Vecchia 2012; Hopkins et al. 2014). The time-step may be further restricted depending on the acceleration of particles, with $\Delta t_{\rm a} = (2\alpha_{\rm k}\epsilon_{\rm grav}/|\boldsymbol{a}|)^{1/2}$ after Power et al. (2003), where $\epsilon_{\rm grav}$ is the force softening length (4 pc) and $\alpha_{\rm k}$=0.02.

The MFM and MFV GIZMO modes use a standard Harten-Lax-van Leer-Contact (HLLC) Riemann solver (Toro 1999; Miyoshi & Kusano 2005) as the default method. In the rare cases where the HLLC solver returns a non-physical result, the code automatically falls back on the slower but more accurate exact solver described in Toro (1997). Flux-limiting is used for the purposes of maintaining numerical stability but we direct the interested reader to appendix B of Hopkins (2015) for a complete and detailed description of how this is implemented in GIZMO.

One of the SPH modes of GIZMO is the traditional density-weighted approach [such as that of GADGET2 (Springel 2005), upon which GIZMO is partly based] which uses a standard artificial viscosity scheme with no additional measures to allow fluid mixing instabilities (e.g. Ritchie & Thomas 2001; Price 2008). While operating as an SPH code, a density estimation is required, the density is determined in the same way as for SPHNG (described in Section 2.2) but with some differences: (i) the typical number of neighbours differs with GIZMO particles having ∼32 rather than ∼58 and (ii) the smoothing length used to scale the kernel is simply h_i rather than the mean h of each particle pair.

Another method available in GIZMO is 'PSPH' (Hopkins 2013; Saitoh & Makino 2013), in which the equations of motion are rearranged to combat fluid mixing instabilities, and also includes artificial conductivity (Price 2008). In the 'traditional' approach, the pressure is calculated using the density estimate, for PSPH the pressure is instead determined from the neighbouring particles using kernel smoothing in the same way as with density.

2.4 Initial conditions and external gravitational potential

We initialize our simulations as a uniform disc with a surface density of 8 $M_\odot$ pc^{-2} and an outer radius of 10 kpc. Gas is distributed vertically with a sech$^2(z/H)$ profile, where the vertical scaleheight, H=0.18 kpc. For our RAMSES runs, we set a density floor of 6.8×10^{-32} g cm^{-3}. The gas is isothermal and has a temperature of 1000 K. We neglect self-gravity in order to investigate the gas response solely to the external gravitational potential. The gas is initially set up with circular orbits according to the underlying gravitational potential. The potential is intended to proxy a rotating stellar mass distribution and is given the logarithmic form from Binney & Tremaine (1987),

$$\psi(r,z)_{\rm disc} = \frac{1}{2}v_0^2\log\left[r^2 + R_{\rm c}^2 + (z/q_\Phi)^2\right] \quad (10)$$

which yields a flat rotation curve with v_0=220 km s^{-1}. The radial and vertical shape of the potential is set by $R_{\rm c}$=1 kpc and q_Φ=0.7. To this potential, we add a spiral perturbation of the form given by Cox & Gómez (2002)

$$\psi(r,\phi,z) = -4\pi G H\rho_0\exp\left(-\frac{r-r_0}{R_{\rm s}}\right) \times\sum_{n=1}^{3}\frac{C_n}{K_nD_n}\cos(n\gamma_{\rm s})\left[\mathrm{sech}\left(\frac{K_nz}{\beta_n}\right)\right]^{\beta_n}, \quad (11)$$

where

$$\gamma_{\rm s} = N\left[\theta + \Omega_{\rm p}t - \frac{r/r_0}{\tan(\alpha)}\right], \quad (12)$$

$$K_n = \frac{nN}{r\sin(\alpha)}, \quad (13)$$

$$\beta_n = K_nH(1+0.4K_nH), \quad (14)$$

$$D_n = \frac{1+K_nH+0.3(K_nH)^2}{1+0.3K_nH}, \quad (15)$$

$$C_1 = 8\pi/3,\ C_2 = 1/2,\ C_3 = 8\pi/15.$$

The parameters r_0=8 kpc, $R_{\rm s}$=7 kpc and H=0.18 kpc set the scaling of the spiral perturbation in three dimensions, N=2 is the number of spiral arms. The pitch angle is α=15° and the pattern speed is $\Omega_{\rm p}=2\times10^{-8}$ rad yr^{-1}. The strength of the spiral perturbation is $\rho_0 = {\rm m_H}n_{\rm H}$ with $n_{\rm H}$=1 atom cm^{-3}. The effective stellar mass of this potential is $\sim10^{11}$ $M_\odot$.

With these parameters, the corotation radius is just beyond the edge of the disc at around 11 kpc. Thus, the gas rotation speed within the entire disc exceeds that of the spiral perturbation and gas shocks at the trailing edge of the perturbation.

We have made runs using each code but without applying the spiral perturbation to determine if any significant features arise within a uniform disc. In this case, we note some weak concentric rings; these are not related to spurious angular momentum transfer but instead are ripples resulting from the imperfect initial pressure equilibrium. We do not think these ripples have any impact on our results for two reasons: (i) the density contrast of the rings is much weaker than that caused by the spiral potential at all radii and (ii) we have repeated our baseline runs starting with the disc in vertical pressure equilibrium to reduce the impact of the ripples and see no difference in the results.

2.5 Full simulation list

A full list of the simulations used in this work is shown in Table 1 with all the parameters that are varied. Four of the runs represent our 'baseline' models which reflect the choices made in most of the work on isolated galaxy simulations with external potentials: these are, *RAMSES-1*, *SPH-1*, *GIZMO-MFM*, and *GIZMO-MFV*. We selected these runs to represent the typical resolutions found in the literature, 4×10^6 particles and four grid cells per Jeans length

Table 1. Overview of the simulation parameters. The top, middle and bottom sections describe the RAMSES, SPHNG, and GIZMO runs, respectively. Column (1): simulation reference name; column (2): maximum refinement level and corresponding physical size of smallest grid cell in pc; column (3): the number of grid cells or particles at the end of the simulation (250 Myr); column (4): number of cells resolving the Jeans length; columns (5) and (6): refinement parameters based on density gradients (see equation 4); columns (7) and (8): Monaghan & Lattanzio (1985) artificial viscosity parameters; column (9): Notes for each run, e.g. code type, solvers and viscosity schemes.

Name	$\ell_{\rm max}$ (pc)	$N_{\rm el}$	$N_{\rm J}$	C_ρ	f_ρ	$\alpha_{\rm v}$	$\beta_{\rm v}$	Notes
RAMSES-E	14 (3.07)	1.808×10^6	4	–	–	–	–	AMR, exact solver + MonCen slope limiter
RAMSES-1	14 (3.07)	9.971×10^5	4	–	–	–	–	AMR, acoustic solver + MinMod slope limiter
RAMSES-2	14 (3.07)	5.362×10^6	8	–	–	–	–	AMR, acoustic solver + MinMod slope limiter
RAMSES-3	14 (3.07)	2.613×10^7	16	–	–	–	–	AMR, acoustic solver + MinMod slope limiter
RAMSES-4	15 (1.54)	2.637×10^7	16	–	–	–	–	AMR, acoustic solver + MinMod slope limiter
RAMSES-5	12 (12.3)	3.037×10^6	4	1.3	5.3	–	–	AMR, acoustic solver + MinMod slope limiter
RAMSES-6	13 (6.15)	5.156×10^6	4	1.3	5.3	–	–	AMR, acoustic solver + MinMod slope limiter
RAMSES-7	14 (3.07)	7.676×10^6	4	1.3	5.3	–	–	AMR, acoustic solver + MinMod slope limiter
RAMSES-8	14 (3.07)	2.366×10^6	4	1.6	5.3	–	–	AMR, acoustic solver + MinMod slope limiter
RAMSES-9	14 (3.07)	6.651×10^6	4	1.0	5.3	–	–	AMR, acoustic solver + MinMod slope limiter
SPH-1	–	4×10^6	–	–	–	1	2	SPH + Monaghan & Lattanzio (1985) viscosity
SPH-2	–	1×10^6	–	–	–	1	2	SPH + Monaghan & Lattanzio (1985) viscosity
SPH-3	–	8×10^6	–	–	–	1	2	SPH + Monaghan & Lattanzio (1985) viscosity
SPH-4	–	4×10^6	–	–	–	0.05	0.1	SPH + Monaghan & Lattanzio (1985) viscosity
GIZMO-MFM	–	4×10^6	–	–	–	–	–	MFM
GIZMO-MFV	–	4.593×10^6	–	–	–	–	–	MFM
GIZMO-MFV-2	–	9.188×10^6	–	–	–	–	–	MFM
GIZMO-MFV-3	–	1.131×10^6	–	–	–	–	–	MFM
GIZMO-SPH-NS	–	4×10^6	–	–	–	–	–	SPH + no viscosity switch (constant viscosity)
GIZMO-SPH-B	–	4×10^6	–	–	–	–	–	SPH + Balsara (1995) viscosity
GIZMO-SPH-C&D	–	4×10^6	–	–	–	–	–	SPH + Cullen & Dehnen (2010) viscosity
GIZMO-PSPH	-	4×10^6	–	–	–	–	–	PSPH + Cullen & Dehnen (2010) viscosity

for particle and AMR runs, respectively.[3] We begin with these runs and later move on to discuss the effect of resolution on a given code.

Given the difficulty in directly comparing resolutions for grid codes with those in particle codes, we illustrate the comparative size of the simulation elements for our baseline simulations in Fig. 1, showing the length of grid cells (for a RAMSES run) and twice the smoothing length (for the GIZMO and SPHNG) versus density.[4] Our goal here is to compare the simulation codes *as they are used in the literature*, but making sure that anything which can overtly affect the physics of the gas is kept the same, e.g. equation of state. Fig. 1 shows that the resolutions of the baseline runs cover roughly the same range of spatial resolutions except that RAMSES extends to much lower densities than the Lagrangian runs. Fig. 1 also illustrates that both Lagrangian codes follow a very tight relation between density and resolution which is not the case with RAMSES.

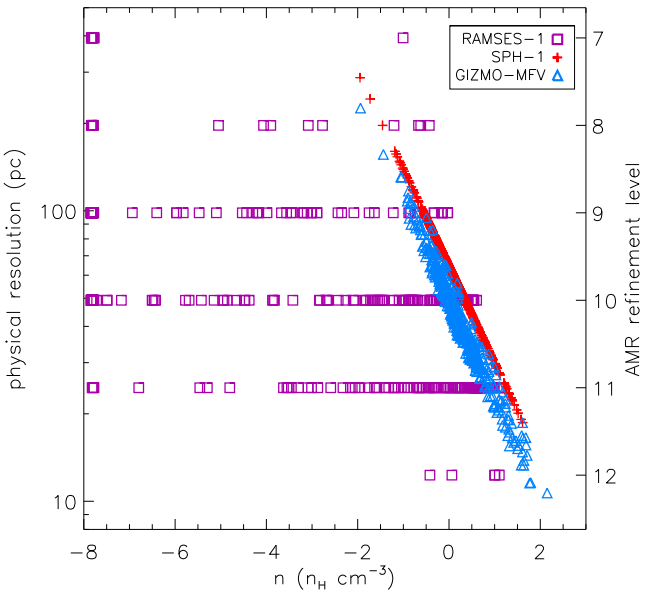

Figure 1. Resolution versus density for the different codes employed in this study. Magenta square symbols are RAMSES leaf cells and the resolution is the length of each cell. For SPHNG (red cross symbols) and GIZMO (blue triangular symbols), the resolution is taken as twice the smoothing length (see Hopkins 2015 for an in-depth discussion on the use of the kernel length in GIZMO). For clarity, we show only 500 randomly selected resolution elements from each of the simulations.

3 RESULTS

All the simulations presented here start from an initially flat surface density distribution within the disc region. In the absence of self-gravity, the interstellar medium (ISM) responds to the external potential and pressure forces leading to an increase in the density of arm regions over time, however we note that we do not form a steady state at any point in our simulations. Initially, there are only two arms which form around the spiral perturbation. This is shortly followed by the development of second pair of arms between the existing ones. These new features branch from the original arms at ∼5 kpc from the centre. The appearance of secondary arm features are well documented as arising from the first ultraharmonic

[3] *RAMSES-1* does not refine the grid to the maximum possible resolution, only reaching level 12 due to the high tolerance of its refinement criteria. All other RAMSES runs make use of the relevant maximum level.

[4] Note that the smoothing length in GIZMO is calculated the same as for an SPH code, but is not used in the same fashion for smoothing the particle distribution.

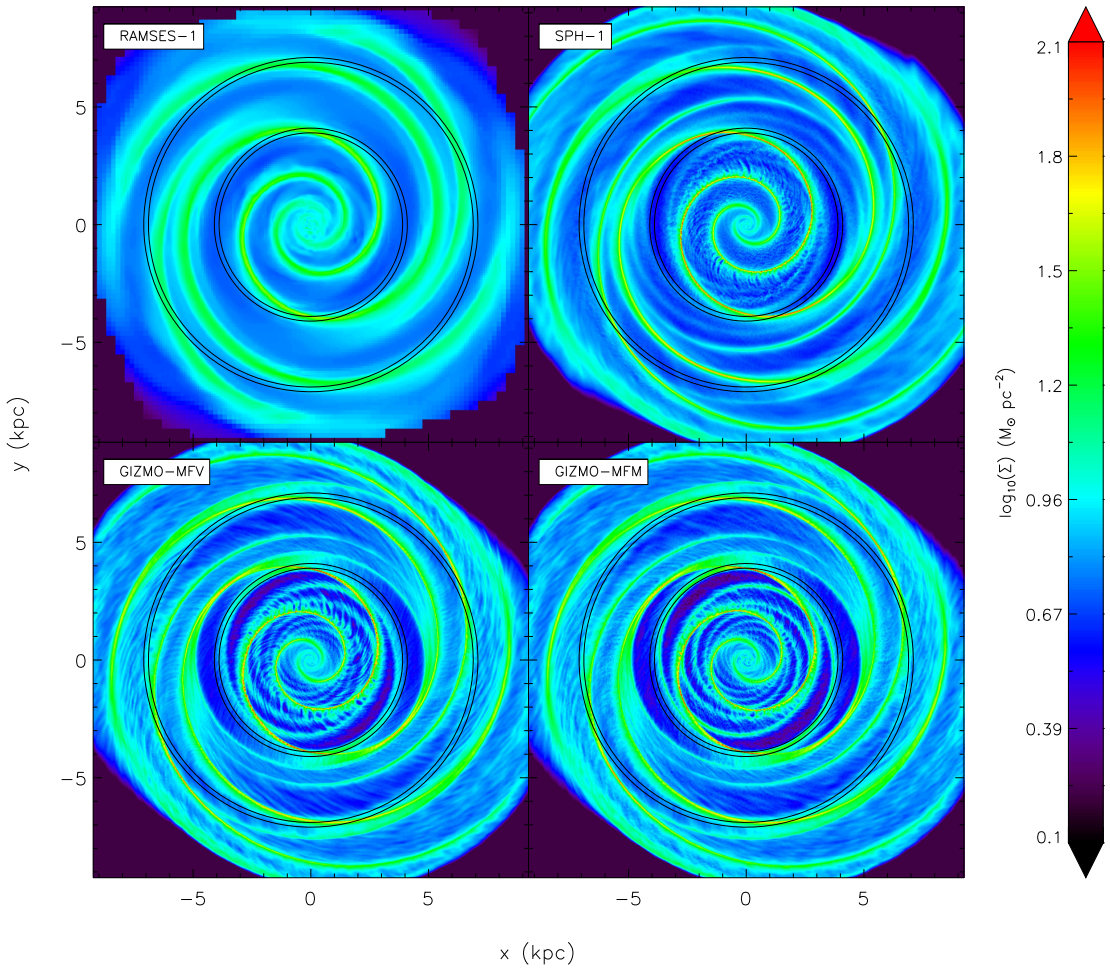

Figure 2. Surface density maps for the baseline models at 250 Myr. The panels show *RAMSES-1* (top-left), *SPH-1* (top-right), *GIZMO-MFV* (bottom-left) and *GIZMO-MFM* (bottom-right) runs. We overlay black circles to indicate the location of the annuli at 4±0.1 and 7±0.1 kpc that are used for the analysis in Section 3.2. SPHNG and GIZMO particles are both smoothed over a cubic spline kernel.

(4:1) resonance (Shu, Milione & Roberts 1973; Patsis, Grosbol & Hiotelis 1997; Chakrabarti, Laughlin & Shu 2003). We show face-on surface density maps for the four baseline models in Fig. 2 in which the original arms and the bifurcating secondary arms are visible. Both features are most clear for *RAMSES-1* in the upper-left panel.

A frequent criticism of grid codes is that angular momentum is not conserved. For the duration of the runs performed in this work (250 Myr), we calculate the loss of angular momentum in the AMR runs as ∼3 per cent and in the SPHNG and GIZMO runs at ∼0.3 per cent. We do not believe that angular momentum loss is a significant cause of the differences we find here, but it may play a more important role in simulations for which self-gravity results in the formation of small-scale eddies.

3.1 Overall structure of the disc with different numerical codes

The different codes tested here respond to the external potential in slightly different ways. Fig. 2 reveals significant differences between the *RAMSES-1* and *SPH-1* runs, with less dense arms and far less fine structure between arms appearing in *RAMSES-1*. The SPHNG and GIZMO runs clearly show the presence of short interarm structures perpendicular to the arms, sometimes referred to as 'spurs' or 'feathers', in the inner part of the disc. These features have been seen in both grid, SPH and AREPO (Smith et al. 2014) simulations, and have been attributed to a number of potential causes including the Kelvin–Helmholz or wiggle instability (Wada & Koda 2004), feathering instability (Lee 2014) and orbit crossing in the spiral arms (Dobbs et al. 2006), although Kim & Ostriker (2006) suggest purely hydrodynamical instabilities disappear or are less evident in 3D. Although there is some slight indication of substructure in the inner part of the *RAMSES-1* disc, distinct spurs are not visible. The differences between the SPHNG and GIZMO runs are largely confined to the strength of the rings found within the central 3 kpc. The two GIZMO runs are not identical: (*GIZMO-MFV* has less coherent rings than *GIZMO-MFM*) but are very similar to one another when compared with the other two runs. The growth of rings in a non-self-gravitating gas disc is also found in grid-based simulations by Shetty & Ostriker (2006), in which *leading* spiral structures develop between the arms near the centre of the disc.

Fig. 3 shows the mean density as a function of distance from the mid-plane for the four baseline simulations at the 250 Myr mark. In this comparison we also include a run, labelled as *RAMSES-E*, which is the same as *RAMSES-1* in all respects except the choice of solver and slope limiter. The majority of our RAMSES simulations use quite a diffusive combination of the acoustic Riemann solver with a MinMod slope limiter, the *RAMSES-E* run uses the exact Riemann

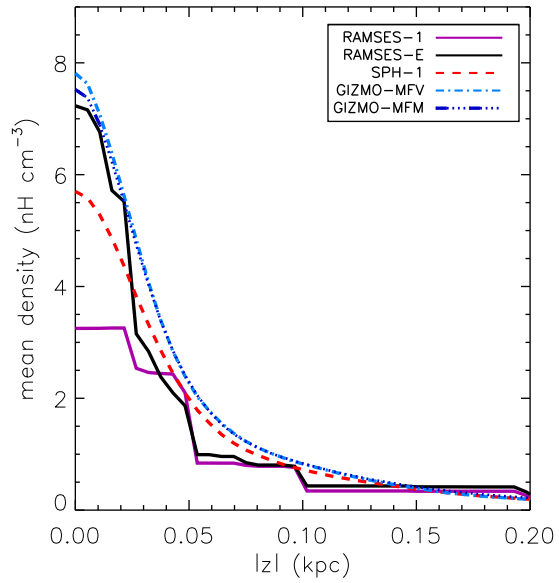

Figure 3. Mean density as a function of distance from the disc plane. Models *RAMSES-1*, *RAMSES-E*, *SPH-1*, *GIZMO-MFV*, and *GIZMO-MFM* are shown as solid magenta, solid black, dashed red, light blue dot–dashed, and dark blue triple dot–dashed lines, respectively.

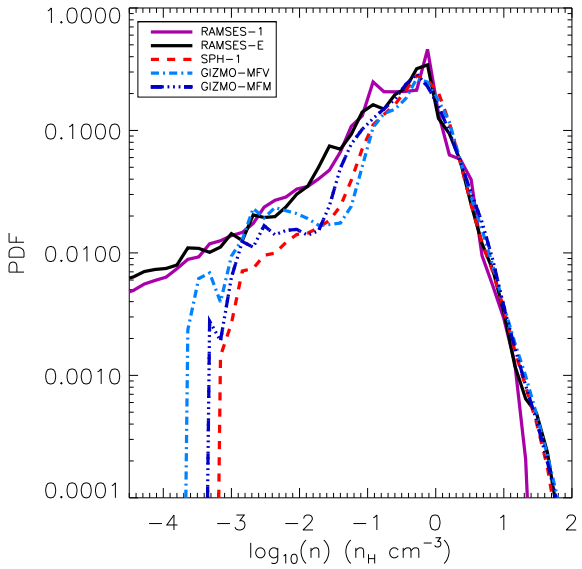

Figure 4. Mass-weighted probability density function for the baseline simulations. Linestyles are the same as in Fig. 3.

solver with the MonCen slope limiter. We use this run to demonstrate the result of using a less diffusive combination. *RAMSES-1* does not reach the same mass concentration as the Lagrangian runs, having only approximately half the mid-plane density of the least concentrated Lagragian run. The *RAMSES-E* run however has a vertical density profile with a slope that is not too dissimilar from the GIZMO runs. The two discs realized with GIZMO are more concentrated than the SPGNG run despite having the same particle resolution.

The one-dimensional structure of the ISM can be examined in the form of a density probability distribution function (PDF) as plotted in Fig. 4. It indicates a consistent density peak for all the runs, but differences do appear in the distributions. The maximum density is curtailed at $10^{1.3}$ n_H cm^{-3} in *RAMSES-1* while the Lagrangian runs extend continuously to around $10^{1.7}$ n_H cm^{-3}. The PDFs of both RAMSES runs extend down to very low densities in contrast with the Lagrangian simulations, simply because there are resolution elements in the RAMSES runs that are not represented by particles in the other codes. We next describe the evolution of the spiral arms over time and will discuss the difference between the codes further in Section 3.3.

3.2 Growth of arm densities

We now consider the surface density of the gas within two annuli, analysing the azimuthal profile at a number of snapshots throughout the runs. The time evolution of azimuthal variations in surface density for the *SPH-1* run is shown in Fig. 5 for annuli at 4±0.1 and 7±0.1 kpc. We show this run as an example which reflects the typical behaviour of all the runs present here. The 4 kpc annulus (upper panel in Fig. 5) illustrates the increasing surface density of the arms until ∼150 Myr, whereupon smaller variations in the peak surface density are present for the remainder of the simulation. The annulus at 7 kpc (lower panel in Fig. 5) shows the same initial increase but with larger variations in the peak arm surface density throughout the run as well as exhibiting the later formation of a secondary arm feature not seen at smaller radii. The effect of the delay between the development of the first and second set of arms is to create oscillations in the maximum surface density. The secondary arms grow downstream in the gas flow and as its density increases, the density of the earlier arm dwindles. This temporarily reduces the maximum surface density until the second arm becomes dominant and the maximum surface density increases again, now representing the secondary arms.

Oscillations in the density of the arms are also expected due to the abrupt activation of the spiral potential. In Woodward (1975), the steepening of spiral waves under the influence of an arm potential that grows over different time-scales is examined, finding that there is an initial period of density oscillation which decays over time. Woodward (1975) also demonstrate that the more extreme and persistent oscillations occur in the runs where the spiral potential is activated over shorter time-scales. A gradually introduced potential reduces the problem of these initial oscillations but for simplicity, and because it is not common practice to do so, we have not employed such a measure. A gradually introduced potential does occur naturally in simulations with a live stellar component without initial spiral structure.

Fig. 6 shows two annuli at 4±0.1 and 7±0.1 kpc and the maximum surface density within each annulus as a function of time for each of our baseline models. Fig. 6 thus illustrates the growth of the density of the arms over time. We include only one of the GIZMO runs here (*GIZMO-MFV*) because its counterpart (*GIZMO-MFM*) is extremely similar, see Fig. 2. There are substantial differences in the surface density of the arms attained for each of the codes. In Fig. 6, we see that particularly at smaller radii the SPHNG runs differ from the GIZMO runs in the timing and amplitude of the oscillations in arm surface density. Lastly we note the considerable difference between *RAMSES-1* and the SPHNG and GIZMO runs, with much lower densities found in the arm regions and lower amplitude oscillations, indeed these oscillations are virtually absent in the *RAMSES-1* run. When one considers *RAMSES-E* there is an interesting difference in the evolution at different radii. At 4 kpc, the oscillations are now to some degree apparent and the arm surface density is only slightly lower than the runs produced with the other codes, however at 7 kpc *RAMSES-E* is not very different to *RAMSES-1* apart from a slight increase in the surface density.

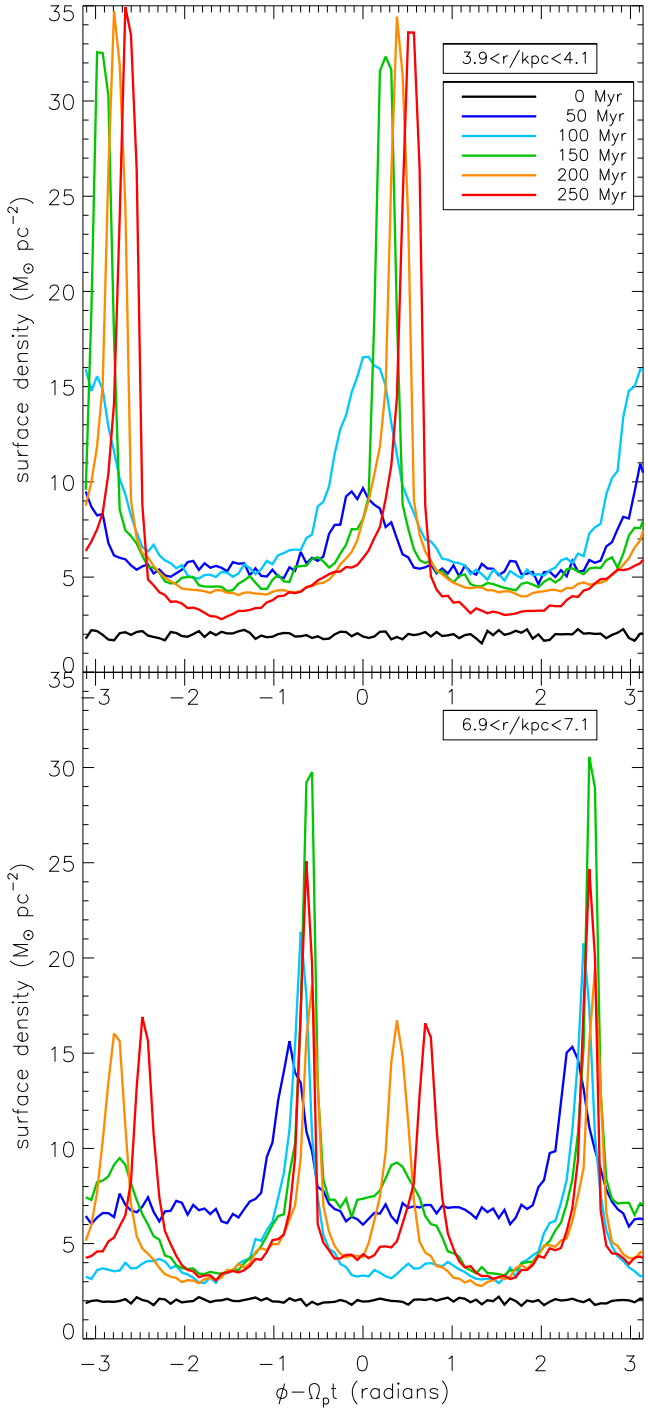

Figure 5. Azimuthal profiles of surface density for the *SPH-1* run in the rotating frame of reference of the external potential. The two panels show different radial annuli (upper: 3.9<*r*/kpc<4.1 and lower: 6.9<*r*/kpc<7.1) and each line represents a different time through the run. We choose the two annular rings as one example at a smaller radius where no secondary arm emerges and a larger radius where it does.

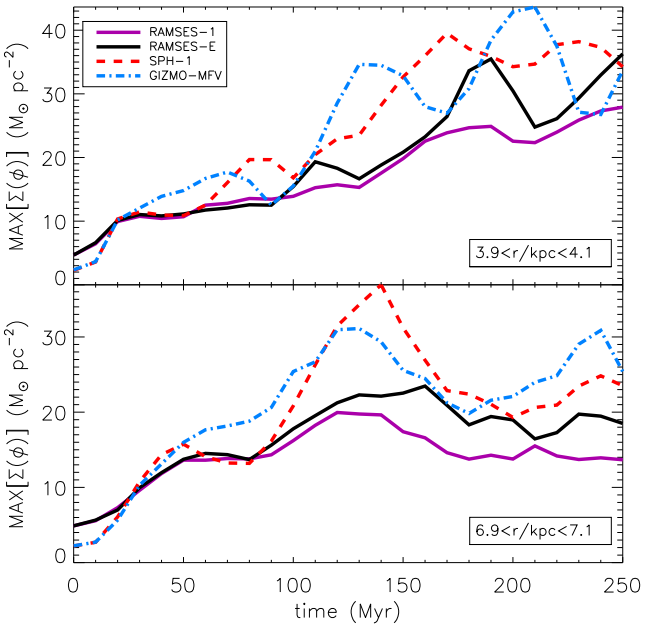

Figure 6. Maximum arm surface density for two annuli (4±0.1 and 7±0.1 kpc) as a function of time. Linestyles are the same as in Fig. 3.

3.3 Dependence of disc evolution on numerical code

Our comparisons of the baseline galaxy disc simulations highlight a number of differences between the codes, particularly the SPHNG and GIZMO runs compared with RAMSES. The maximum density in the spiral arms, maximum density in the mid-plane and the degree of interarm structure including the secondary branches are greater with GIZMO and SPHNG, more so in the GIZMO runs. We now consider which characteristics of the codes might lead to these differences. The most obvious possibilities are that the codes reach different effective resolutions, or the inclusion of artificial viscosity in SPHNG. We discuss viscosity next, and describe resolution tests of the codes in Sections 3.4 and 3.5.

SPHNG runs employ artificial viscosity to allow shock capturing, but artificial viscosity is not required in either RAMSES or GIZMO as both use a Riemann solver. For this reason, we do not believe that the discrepancy between *RAMSES-1* and the other runs in Fig. 6 is due to different viscosity schemes. Additionally we have run SPHNG with reduced viscosity parameters, α_v=0.05 and β_v=0.1. These viscosity parameters yield an arm surface density growth rate (without oscillations) similar to *RAMSES-1* but shock capturing is compromised by reducing the artificial viscosity so harshly.

GIZMO can also be operated as an SPH code with various viscosity switches. In order to further eliminate the influence of viscosity as the cause of differences between the inviscous MFM and MFV methods of GIZMO and the artificially viscous SPHNG, we have run our model with a number of these modes (detailed in Table 1).

These models are compared with *GIZMO-MFV* and *SPH-1* in Fig. 7. It is clear that the artificial viscosity scheme makes little difference and all GIZMO modes behave in roughly the same way. Therefore, the cause of the discrepancy found between the SPHNG and GIZMO runs is not a basic difference in the SPH or the meshless Lagrangian methods of Hopkins (2015). The fact that, despite some minor differences, the evolution of the GIZMO SPH runs are far more similar to *GIZMO-MFV* than they are to *SPH-1* means that it is likely that some aspect of the code is responsible for the offset between the two codes that is separate from the fundamental methodologies. We have not explored the codes in sufficient detail to offer a definitive explanation of this but note that, apart from the core hydrodynamics solver, all components of GIZMO are commons to both the MFV/MFM and SPH modes. The different results found with SPHNG and GIZMO may therefore be due to differences in the neighbour finding process, the time-step criteria or time integration.

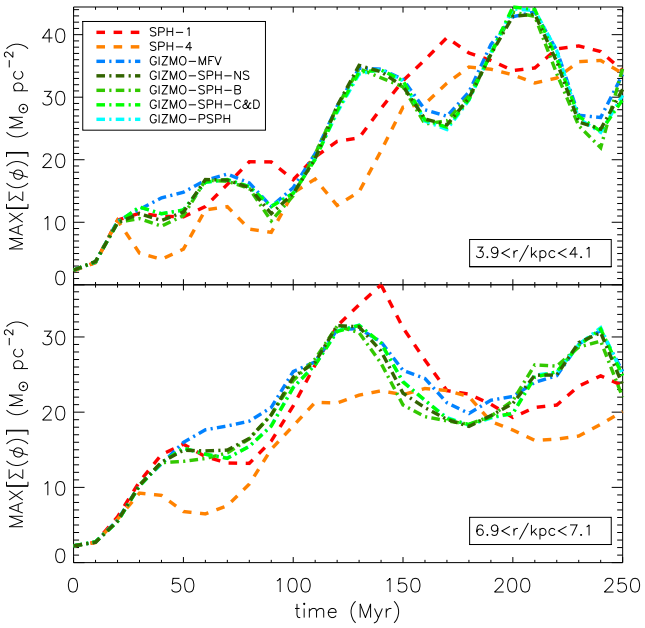

Figure 7. Maximum arm surface density for two annuli (4±0.1 and 7±0.1 kpc) as a function of time for various SPH modes run with GIZMO and the baseline GIZMO run, *GIZMO-MFV*. We also show *SPH-1* and *SPH-4* to compare the evolution with lower artificial viscosity. Red and orange dashed lines are SPHNG, *GIZMO-MFV* is show in blue, and runs and green–cyan lines are runs with GIZMO in SPH mode. The *GIZMO-PSPH* almost exactly follows the *GIZMO-SPH-C&D* run.

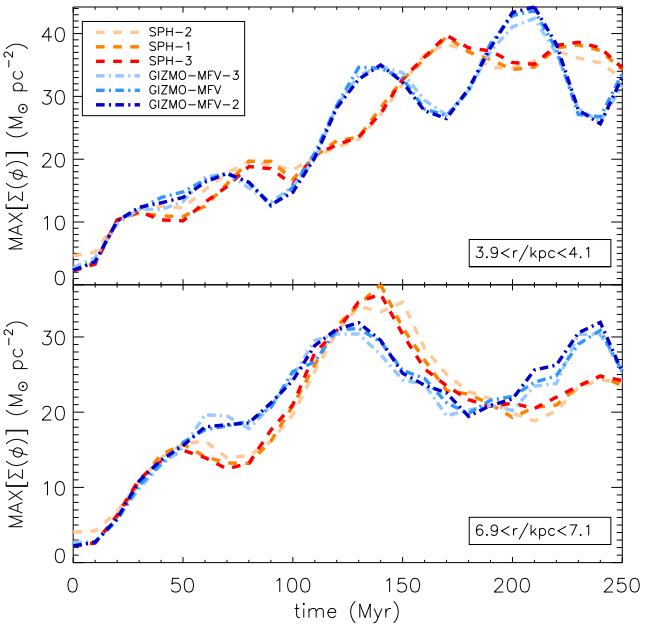

Figure 8. Maximum arm surface density as a function of time for SPHNG (red/orange dashed lines) and GIZMO (blue dot–dashed lines) runs with different resolutions (initially with 1, 4 and 8 million particles). Darker colours represent greater resolutions.

We note that the SPH kernel for GIZMO and SPHNG is a cubic spline but that the smoothing length is defined differently with the smoothing lengths in GIZMO being smaller (see Fig. 1). Exploring this, we have rerun *GIZMO-MFM* and *GIZMO-SPH-C&D* with the same particle smoothing lengths as our SPHNG runs and find that it makes very little difference to the arm surface density and density PDF, however the vertical concentration is slightly reduced and therefore closer to the profile of *SPH-1*. We also observe a slight weakening of interarm structures when using larger smoothing lengths.

3.4 Resolution of SPHNG and GIZMO runs

In this section, we consider how our two Lagrangian codes (SPHNG and GIZMO) evolve differently with varying mass resolution. The evolution of the arm surface density for runs that initially have 1×10^6, 4×10^6, and 8×10^6 particles are shown in Fig. 8. The number of particles in SPHNG and GIZMO MFM runs is fixed, but the GIZMO MFV runs allow particle splitting and increase the number of particles over the course of the simulation. The final number of particles for *GIZMO-MFV-3*, *GIZMO-MFV*, and *GIZMO-MFV-2* are 1.131×10^6, 4.593×10^6, and 9.188×10^6, respectively.

For both of the Lagrangian codes, the evolution of the peak surface density is invariant with resolution, i.e. the rate of arm growth is unaffected. SPHNG and GIZMO runs do not show much difference in the surface density of the galaxy arms as a function of resolution, but we do find differences between the codes themselves. Despite the codes presenting a similar time-averaged growth curve, the oscillations discussed in the previous section are offset in time (see Fig. 8).

Despite the invariance of the maximum surface density with resolution, we do see that the vertical density profiles are steeper for runs with higher resolution, as shown in Fig. 9 and consequently the peak volume density is enhanced by resolution. As the particle resolution is increased (and the smoothing length shortens), the vertical density profiles is improved. The GIZMO runs exhibit steeper density profiles than do the SPHNG, this is partially due to the different definition and use of the smoothing length/particle domain.

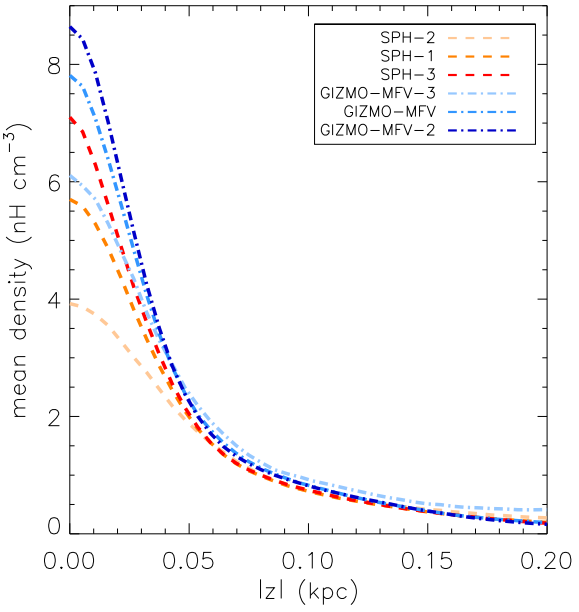

Figure 9. The vertical density profiles of SPHNG (red/orange dashed lines) and GIZMO (blue dot–dashed lines) runs with different resolutions (initially with 1, 4 and 8 million particles). Darker colours representing greater resolutions.

The mass-weighted density PDFs for our resolution comparison of Lagrangian runs are shown in Fig. 10. For both SPHNG and GIZMO, increased resolution broadens the distribution and increases

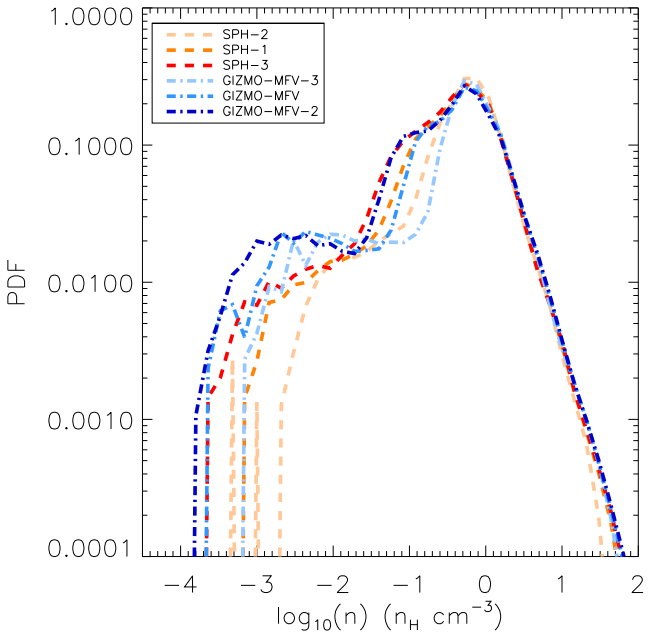

Figure 10. Mass-weighted probability density function for Lagrangian simulations with different resolutions. Lines show SPHNG (red/orange dashed lines) and GIZMO (blue dot–dashed lines) runs that initially have 1, 4 and 8 million particles, darker colours represent greater resolutions.

the fraction of gas at lower densities. GIZMO runs have a consistently narrower distribution in the high-density peak but extend to lower densities than their SPHNG counterparts. The maximum density achieved by all the runs is quite consistent with the sole exception of *SPH-2* (the lowest resolution SPH run) which is truncated around 0.2 dex below the others.

3.5 Resolution of AMR runs

Resolution in AMR simulations is not a linear characteristic. We can change the minimum and maximum refinement levels or the parameters governing refinement. In this section, we compare a number of approaches to varying the resolution within RAMSES. In addition to the previously shown run (*RAMSES-1*) which employs a Jeans length refinement criterion with the typical threshold N_J=4, we now test the effect of varying N_J and C_ρ, which control grid refinement according to the local Jeans length and density gradients, respectively. We also vary ℓ_{max} which is the upper limit on the grid level. The specific parameters used in each run are detailed in Table 1. We now discuss these three parameters that control refinement in turn using the mass-weighted density PDFs in Fig. 11 and the maximum surface density within annuli at 4±0.1 and 7±0.1 kpc as a function of time in Fig. 12.

First, we consider the effect of increasing ℓ_{max} which permits the code to refine the grid to higher levels. For this, we direct the reader to the runs shown in the top panel of Fig. 11, i.e. we compare *RAMSES-3* and *RAMSES-4* which have ℓ_{max}= 14 and 15, respectively, but with all other refinement criteria the same. We also compare three runs that have an alternative set of refinement criteria to the previous two which use ℓ_{max}= 12, 13 and 14 (*RAMSES-5*, *RAMSES-6* and *RAMSES-7*), see Table 1 for details. We see here that increasing ℓ_{max} does not have an enhancing influence on these simulations, indeed *RAMSES-3* and *RAMSES-4* have identical density PDFs. We further examine the growth of the spiral arm surface density in the left-hand panels of Fig. 12. Again we note that varying

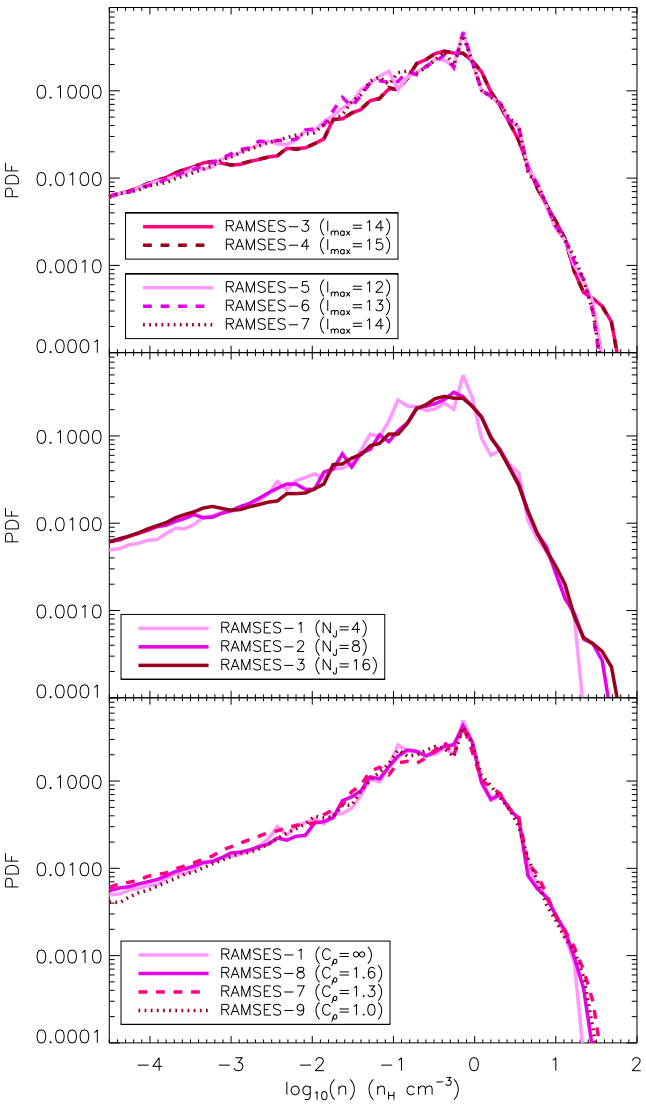

Figure 11. Mass-weighted probability density functions for RAMSES runs with different refinement schemes. In the top panel, we compare runs with differing values of ℓ_{max}. The middle panel compares different N_J values. The lower panel shows runs with different density gradient refinement thresholds (C_ρ). Darker colours represent greater resolutions: see Table 1 for details of the parameters used in each run.

ℓ_{max} makes almost no difference to the evolution of the galactic disc, although in this case there is a marginal reduction in arm surface density as ℓ_{max} increases for both the 4 and 7 kpc annuli in the case of *RAMSES-5*, *RAMSES-6* and *RAMSES-7*.

We now consider whether increasing the number of grid cells that resolve the Jeans length (N_J) has an influence on the simulations. *RAMSES-1*, *RAMSES-2*, and *RAMSES-3* have N_J = 4, 8 and 16, respectively, and are shown in the middle panel of Fig. 11. We note a marked increase in the maximum density achieved when N_J takes greater values. In fact, the densest end of the density distribution function for *RAMSES-3* (for which N_J=16) is consistent with that found in the GIZMO and SPHNG runs. The central panels of Fig. 12 illustrate how increasing N_J alters the growth of the spiral arms. We see that greater values of N_J produce higher surface densities at 4 kpc. At 7 kpc we see simply that as N_J increases, the oscillations that are clearly present in SPHNG and GIZMO runs (see Fig. 5) become more apparent.

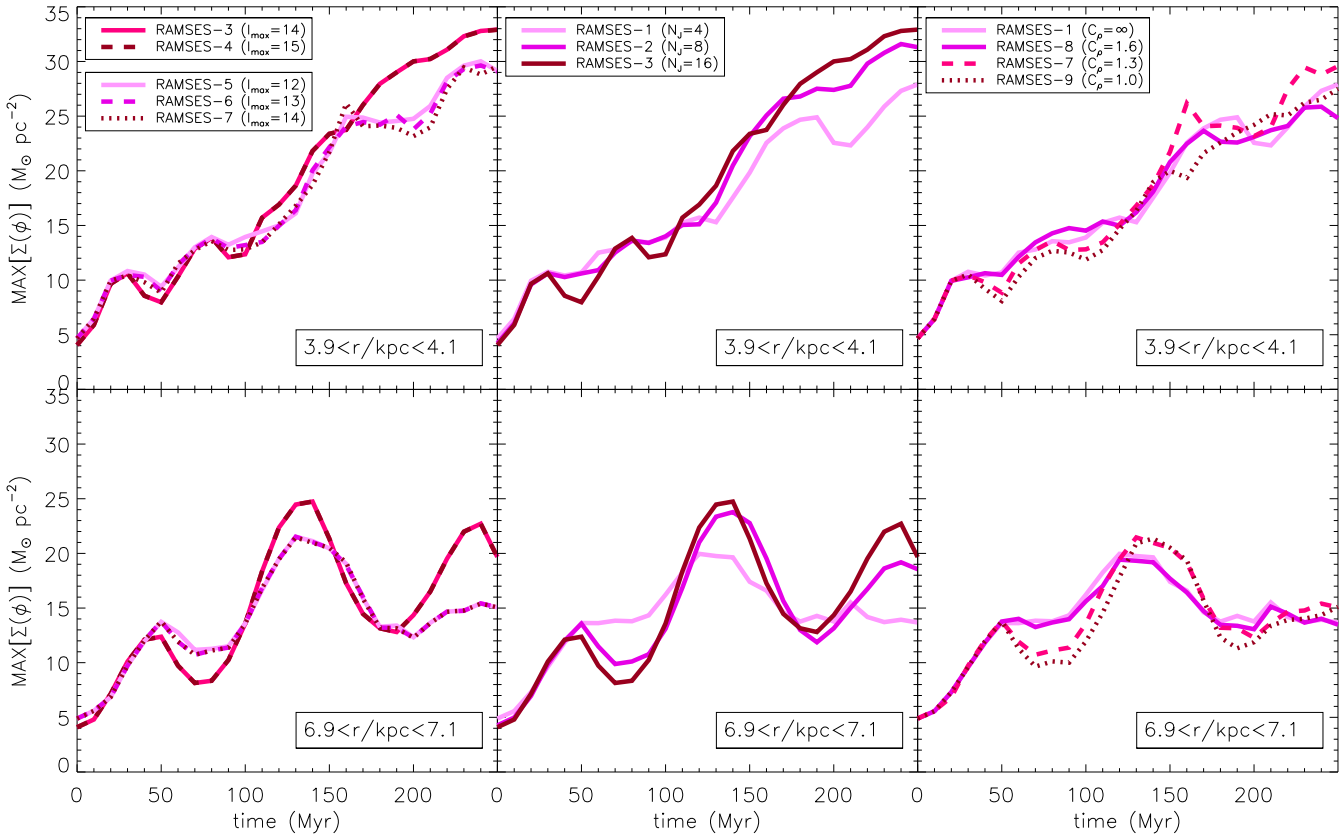

Figure 12. Maximum arm surface density as a function of time for RAMSES runs with various refinement schemes. The upper panels illustrate the evolution of the spiral arms at 4±0.1 kpc and the lower panels at 7±0.1 kpc. The left-hand panels compare runs with differing values of $\ell_{\rm max}$. The middle panels show runs with different $N_{\rm J}$. The right-hand panels show runs with different density gradient refinement thresholds (C_ρ). Darker colours represent greater resolutions: see Table 1 for details of the parameters used in each run.

One key strength of grid codes is the ability to resolve sharp density contrasts, achieved partially through the use of refinement criteria based on the local gradient in hydrodynamical variables. The final comparison we make varies the threshold that controls grid refinement based on the density gradients, C_ρ. A lower value of C_ρ corresponds, in principle, to greater resolution. We compare *RAMSES-8*, *RAMSES-7*, and *RAMSES-9* (in order of decreasing C_ρ) in the bottom panel of Fig. 11 and right-hand panel of Fig. 12. In this comparison, we also include *RAMSES-1* which effectively has an infinite threshold, i.e. no grid refinement is permitted based on density gradients. We recall here that *RAMSES-1* does not make use of grid levels 13 and 14 and choosing a finite C_ρ allows the adaptive grid to make use of these levels.

In Fig. 11 (lower panel), we note a slight increase in the maximum density value as C_ρ decreases. The lowest C_ρ run (*RAMSES-9*) does not follow the trend of increasing peak density but this is likely linked to it having fewer grid cell than *RAMSES-7* despite its lower C_ρ. We therefore note that the higher density correlates more with the number of grid cells than with the gradient refinement threshold and likely does not reflect the ability of this refinement scheme to place cells in appropriate locations. Examining the inner annulus in Fig. 12 (upper right panel) gives no clear indication of whether the value of C_ρ has any impact on the arm surface density. The outer annulus (lower right panel in Fig. 12) however suggests that the lower threshold does enhance the oscillations of the arm surface density.

We find that, in the context of our galactic disc simulations, the grid structure is very sensitive to C_ρ. We therefore find that decreasing C_ρ can lead to very little increase in resolution (because the grid is already refined to an extent by the Jeans length criteria), or it can refine the grid to such a degree that the simulation time increases disproportionately compared with other approaches. We find that despite the lower value of C_ρ in *RAMSES-9* than in *RAMSES-7* (which should mean that it refines grid cells more easily), we actually have fewer grid cells by the end of the simulation. For the first 150 Myr, the number of grid cells found in *RAMSES-9* is much higher than in *RAMSES-7*, consistent with its lower refinement threshold, but it then declines gradually to the value found in Table 1.

Surface density maps for our baseline RAMSES run and the two resolution tests that are most similar to the Lagrangian runs (*RAMSES-3* and *RAMSES-9*) are compared in Fig. 13. Both these higher resolution runs have sharper arms but also exhibit more interarm structure. *RAMSES-3*, which resolves the Jean's length with 16 grid cells, presents hints of the rings seen in the Lagrangian runs and secondary arms that extend further beyond the branching point (as seen in the Lagrangian runs in Fig. 2). Similar structures are seen in *RAMSES-9*, although more feathers are seen near the centre and the secondary arms are not as extended. The emergence of these structures in the higher resolution runs hints that concordance with the Lagrangian runs is closer.

4 DISCUSSION AND CONCLUSIONS

We have examined how an isothermal gas disc evolves under the influence of an external spiral potential when realized with different hydrodynamical methods (RAMSES, SPHNG, and GIZMO) and as a

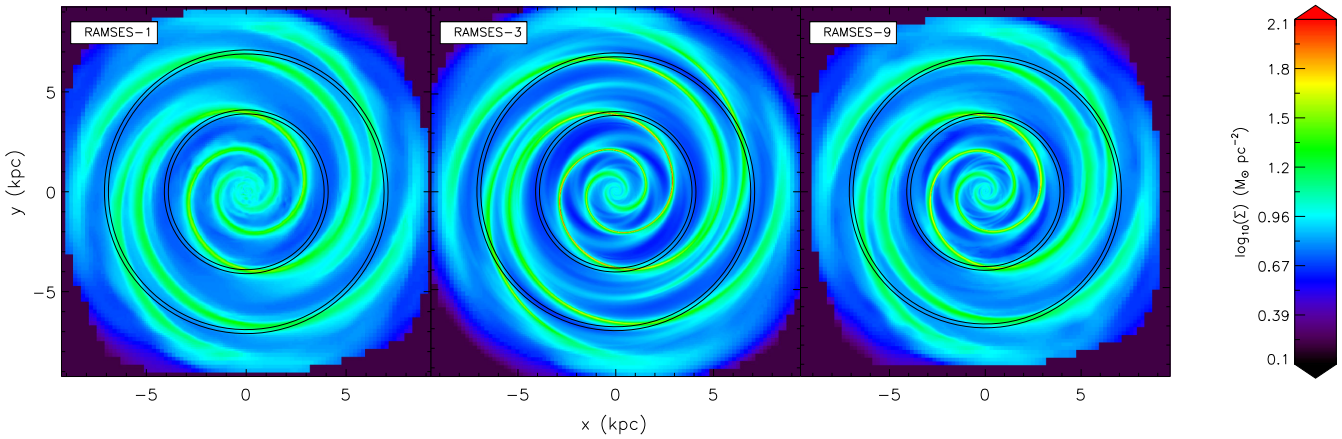

Figure 13. Surface density maps for our baseline RAMSES run (*RAMSES-1*) and for the two highest resolution runs using N_J and C_ρ criteria (*RAMSES-3*, and *RAMSES-9*, respectively) after 250 Myr have passed. Black circles indicate the location of the annuli at 4±0.1 and 7±0.1 kpc used in our analysis. The most significant difference between these runs is the concentration of the arms and the presence of more interarm structure in the higher resolution runs.

function of resolution. With similar resolutions to those found in the literature (and using the 'acoustic' solver with a MinMod slope limiter), we find that our AMR code, RAMSES, generates a weaker density contrast between the arm and interarm region, less steep vertical profiles and lower arm surface densities than we see when using the Lagrangian codes (SPHNG and GIZMO) as well as less interarm structure. When additional refinement measures are used, RAMSES generates results in better agreement with the other codes. If a less diffusive set up is used (i.e. an 'exact' solver with a MonCen slope limiter), then a measure of similarity is also achieved in the resolution of physical structures, vertical density profile, and the density PDF. The growth of spiral arm surface density is also enhanced to a similar level as seen in the Lagrangian codes but only at smaller radii: arm surface density is still relatively low at greater radii. In all codes, we also see oscillations in the peak arm surface density, which appear to be associated with the development of secondary arms, but the oscillations tend to be very weak with the baseline RAMSES model. The GIZMO and SPHNG codes also display small differences, namely that GIZMO produces the highest densities and surface densities.

To test why the differences occurred between the codes, we considered the resolution and viscosity. We found resolution had little effect on the Lagrangian runs, except for an increasing steepness of the vertical density profiles with resolution. Whilst artificial viscosity could potentially affect the results with the SPHNG code, our fiducial GIZMO runs do not include artificial viscosity, so this cannot explain the discrepancies between GIZMO and RAMSES. Furthermore, using GIZMO in 'SPH mode' (with constant, Balsara 1995 and Cullen & Dehnen 2010 viscosity schemes and with pressure-entropy based SPH) we find very little difference between those runs and the equivalent MFM and MFV GIZMO modes (without viscosity). By reducing the artificial viscosity by a factor of 20, the SPHNG models produced more similar results to the baseline RAMSES run, however this dramatic change in viscosity greatly reduces the ability of SPHNG to accurately capture shocks. Thus we believe that the differences between RAMSES and the SPHNG/GIZMO runs, and between SPHNG and GIZMO are not due to viscosity. Differences in the choice of smoothing length and the functional form of the kernel partly account for the differences between our SPHNG and GIZMO runs.

We also investigated varying the resolution in the RAMSES code using a number of approaches. We first used the most intuitive approach, increasing the maximum refinement level $\ell_{\rm max}$, but surprisingly this made almost no difference. The reason for this was because the code was simply not saturating the maximum refinement level. We stress here though that our simulations are isothermal and do not include self-gravity, gas cooling, or any sources of forcing other than our external potential. These processes could drive the density up and force the grid to refine even without changing the refinement criteria, though it is not clear that this would improve the modelling of spiral arms more generally or, for example, the initial development of Jeans' instabilities.

We secondly tried using refinement criteria based on density gradients (the kind conventionally employed in idealized code comparison tests), which we might expect would lead to greater refinement and better agreement with the arm densities seen with the other codes. However, we find that the parameters controlling such refinement schemes are not trivial to select a priori. We tend to find only slight differences in the density PDF and, observe a small increase in the scale of oscillations in the arm surface density over time. Our model *RAMSES-9* shows best agreement with the other codes, but we note that in this run we were refining a significant fraction of the disc, which leads to a large increase in computing time due to refining regions unnecessarily. Again, we note that these results are true for our particular model choices and hydrodynamical gradient criteria may be of benefit under other frameworks, e.g. when supernovae blastwaves are included.

Thirdly, we examined increasing refinement by increasing the number of grid cells resolving the Jeans length, N_J. We find that the greatest, and most reliable, improvement (in the sense of providing more concordance with other methods) in the density PDF, vertical density profile and evolution of the surface density is found by increasing N_J. By changing N_J from 4 to 16, we find greater similarities between our AMR and Lagrangian tests. In addition to the spiral arms, the resolution of the interarm regions is also important, one reason being that dense structures leaving the refined region should be preserved. To some extent, our highest resolution RAMSES runs do preserve the interarm structures found in the Lagrangian runs. Although there are still differences between RAMSES and the other codes, this finding is in line with other comparison tests that show more similar results tend to be attained with grid and particle-based codes (Tasker et al. 2008; Kitsionas et al. 2009; Price & Federrath 2010) if the resolution is comparable or greater in the grid code. Our results also seem to concur with the idea in Price & Federrath

(2010) that grid-based codes tend to be less effective at resolving denser regions of simulations.

Our findings that the density in spiral arms differs according to different numerical codes (we stress that here we mean 'code' and not 'method'), and further that the growth of the maximum density differs, may have implications for studying star formation in spiral arms. The different densities may impact the time-scales for gravitational collapse, properties and number of molecular clouds, and the rate and efficiency of star formation. This may be even more relevant for simulations with transient spiral arms, whereby the arms come and go with time, and the time for gas to accumulate into dense structures in the spiral arms may be fundamentally limited. We have demonstrated that by refining adaptive grid simulations further than is usual, similar arm density growth rates are achieved as with Lagrangian codes. We note that a quasi-Lagrangian refinement scheme could also be applied in the case where transient spiral arms arise due to a live stellar disc, but we do not test this idea in this work. We have also not examined how refinement changes once self-gravity is included, but likewise leave this to future studies.

To conclude, our key result is that caution is required when justifying the use of four grid cells per Jeans' length as a resolution scheme with AMR (or static grid) simulations. This condition is necessary but may not be sufficient depending on the hydrodynamical solver. For the simulations performed here, the Jeans length needs to be resolved by at least 16 grid cells to achieve a similar result to that found with Lagrangian codes. The authors are aware of only one study in the field of isolated galaxy simulations where the refinement criteria is set to a higher value, Petit et al. (2015), wherein 32 cells are used (note the authors also applied a quasi-Lagrangian refinement scheme). We find some dependence on the solver (in particular using a less diffusive solver also produces better resolution), and in general the criteria for the resolution may depend on the exact nature of the code. But given that more diffusive set ups are likely frequently used, we believe that this result is an important one. We find that increasing $N_{\rm J}$ appears to be the most effective and simplest means of increasing the resolution in regions of interest in a galactic disc, in particular to study processes such as spiral shocks and molecular cloud formation.

ACKNOWLEDGEMENTS

CGF, CLD and LK acknowledges funding from the European Research Council for the FP7 ERC starting grant project LOCALSTAR. We thank the anonymous referee for a very constructive report. This work used the DiRAC Complexity system, operated by the University of Leicester IT Services, which forms part of the STFC DiRAC HPC Facility (www.dirac.ac.uk). This equipment is funded by BIS National E-Infrastructure capital grant ST/K000373/1 and STFC DiRAC Operations grant ST/K0003259/1. DiRAC is part of the National E-Infrastructure. Calculations were also performed on Cray XC30 at the Center for Computational Astrophysics, National Astronomical Observatory of Japan.

REFERENCES

Agertz O. et al., 2007, MNRAS, 380, 963
Agertz O., Romeo A. B., Grisdale K., 2015, MNRAS, 449, 2156
Balsara D. S., 1995, J. Comput. Phys., 121, 357
Bate M. R., Bonnell I. A., Price N. M., 1995, MNRAS, 277, 362
Benz W., 1990, in Buchler J. R., ed., Numerical Modelling of Nonlinear Stellar Pulsations Problems and Prospects. Kluwer, Dordrecht, p. 269
Berger M. J., Colella P., 1989, J. Comput. Phys., 82, 64
Berger M. J., Oliger J., 1984, J. Comput. Phys., 53, 484
Binney J., Tremaine S., 1987, Galactic dynamics. Princeton Univ. Press, Princeton, NJ
Bournaud F., Elmegreen B. G., Teyssier R., Block D. L., Puerari I., 2010, MNRAS, 409, 1088
Chakrabarti S., Laughlin G., Shu F. H., 2003, ApJ, 596, 220
Cha S.-H., Whitworth A. P., 2003, MNRAS, 340, 73
Cox D. P., Gómez G. C., 2002, ApJS, 142, 261
Cullen L., Dehnen W., 2010, MNRAS, 408, 669
Dobbs C. L., Pringle J. E., 2013, MNRAS, 432, 653
Dobbs C. L., Bonnell I. A., Pringle J. E., 2006, MNRAS, 371, 1663
Dobbs C. L., Glover S. C. O., Clark P. C., Klessen R. S., 2008, MNRAS, 389, 1097
Dobbs C. L. et al., 2014, in Henrik Beuther H., Klessen R. S., Dullemond C. P., Henning T., eds, Protostars and Planets VI. Univ. Arizona Press, Tucson, AZ, p. 3
Durier F., Dalla Vecchia C., 2012, MNRAS, 419, 465
Fehlberg E., 1969, NASA technical report, R-315, Low-order Classical Runge–Kutta Formulas with Stepsize Control and Their Application to Some Heat Transfer Problems. Washington, US
Frenk C. S. et al., 1999, ApJ, 525, 554
Fujimoto Y., Tasker E. J., Wakayama M., Habe A., 2014, MNRAS, 439, 936
Gaburov E., Nitadori K., 2011, MNRAS, 414, 129
Grand R. J. J., Kawata D., Cropper M., 2012, MNRAS, 421, 1529
Hopkins P. F., 2013, MNRAS, 428, 2840
Hopkins P. F., 2015, MNRAS, 450, 53
Hopkins P. F., Kereš D., Murray N., Quataert E., Hernquist L., 2012, MNRAS, 427, 968
Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, MNRAS, 445, 581
Inutsuka S.-I., 2002, J. Comput. Phys., 179, 238
Kereš D., Vogelsberger M., Sijacki D., Springel V., Hernquist L., 2012, MNRAS, 425, 2027
Khokhlov A., 1998, J. Comput. Phys., 143, 519
Khoperskov S. A., Vasiliev E. O., Sobolev A. M., Khoperskov A. V., 2013, MNRAS, 428, 2311
Kim J.-h. et al., 2014, ApJS, 210, 14
Kim W.-T., Ostriker E. C., 2002, ApJ, 570, 132
Kim W.-T., Ostriker E. C., 2006, ApJ, 646, 213
Kitsionas S. et al., 2009, A&A, 508, 541
Klessen R. S., Heitsch F., Mac Low M.-M., 2000, ApJ, 535, 887
Kritsuk A. G. et al., 2011, ApJ, 737, 13
Lanson N., Vila J. P., 2008a, SIAM J. Numer. Anal., 46, 1912
Lanson N., Vila J. P., 2008b, SIAM J. Numer. Anal., 46, 1935
Lee W., 2014, ApJ, 792, 122
Mata-Chávez M. D., Gómez G. C., Puerari I., 2014, MNRAS, 444, 3756
Miyoshi T., Kusano K., 2005, J. Comput. Phys., 208, 315
Monaghan J. J., 1992, ARA&A, 30, 543
Monaghan J. J., 1997, J. Comput. Phys., 136, 298
Monaghan J. J., Lattanzio J. C., 1985, A&A, 149, 135
O'Shea B. W., Nagamine K., Springel V., Hernquist L., Norman M. L., 2005, ApJS, 160, 1
Pakmor R., Springel V., 2013, MNRAS, 432, 176
Patsis P. A., Grosbol P., Hiotelis N., 1997, A&A, 323, 762
Pearce F. R. et al., 1999, ApJ, 521, L99
Petit A. C., Krumholz M. R., Goldbaum N. J., Forbes J. C., 2015, MNRAS, 449, 2588
Power C., Navarro J. F., Jenkins A., Frenk C. S., White S. D. M., Springel V., Stadel J., Quinn T., 2003, MNRAS, 338, 14
Price D. J., 2008, J. Comput. Phys., 227, 10040
Price D. J., Federrath C., 2010, MNRAS, 406, 1659
Price D. J., Monaghan J. J., 2004, MNRAS, 348, 139
Renaud F. et al., 2013, MNRAS, 436, 1836
Ritchie B. W., Thomas P. A., 2001, MNRAS, 323, 743
Robertson B. E., Kravtsov A. V., 2008, ApJ, 680, 1083
Saitoh T. R., Makino J., 2009, ApJ, 697, L99
Saitoh T. R., Makino J., 2013, ApJ, 768, 44

Saitoh T. R., Daisaka H., Kokubo E., Makino J., Okamoto T., Tomisaka K., Wada K., Yoshida N., 2008, PASJ, 60, 667
Scannapieco C. et al., 2012, MNRAS, 423, 1726
Shetty R., Ostriker E. C., 2006, ApJ, 647, 997
Shu F. H., Milione V., Roberts W. W., Jr, 1973, ApJ, 183, 819
Smith R. J., Glover S. C. O., Clark P. C., Klessen R. S., Springel V., 2014, MNRAS, 441, 1628
Springel V., 2005, MNRAS, 364, 1105
Springel V., 2010, MNRAS, 401, 791
Tasker E. J., Tan J. C., 2009, ApJ, 700, 358
Tasker E. J., Brunino R., Mitchell N. L., Michielsen D., Hopton S., Pearce F. R., Bryan G. L., Theuns T., 2008, MNRAS, 390, 1267
Tasker E. J., Wadsley J., Pudritz R., 2015, ApJ, 801, 33
Teyssier R., 2002, A&A, 385, 337
Toro E., 1997, Riemann Solvers and Numerical Methods for Fluid Dynamics: A Practical Introduction. Springer-Verlag, Berlin
Toro E., 1999, Riemann Solvers and Numerical Methods for Fluid Dynamics: A Practical Introduction, 2nd ed. Springer-Verlag, Berlin
Torrey P., Vogelsberger M., Sijacki D., Springel V., Hernquist L., 2012, MNRAS, 427, 2224
Truelove J. K., Klein R. I., McKee C. F., Holliman J. H., II, Howell L. H., Greenough J. A., 1997, ApJ, 489, L179
Wada K., 2008, ApJ, 675, 188
Wada K., Koda J., 2004, MNRAS, 349, 270
Wada K., Norman C. A., 2007, ApJ, 660, 276
Wadsley J. W., Veeravalli G., Couchman H. M. P., 2008, MNRAS, 387, 427
Whitehurst R., 1995, MNRAS, 277, 655
Williamson D. J., Thacker R. J., 2012, MNRAS, 421, 2170
Williamson D. J., Thacker R. J., Wurster J., Gibson B. K., 2014, MNRAS, 442, 3674
Woodward P. R., 1975, ApJ, 195, 61

This paper has been typeset from a TEX/LATEX file prepared by the author.